How not to do research

Peter E. Hodgson[*]

Scientific research is often portrayed as a rather mechanical routine consisting of amassing facts and inducing theories to account for them. The reality is quite different, and we really do not understand how it is done. If we cannot say how to do research, the next best thing is to say how not to do it. It would certainly be presumptuous of me to pretend that I know how to do research.

Everyone who wants to undertake research is faced with the problem of how to get started. To overcome this, beginners are assigned a supervisor, who suggests a problem to tackle and how it might be tackled. This is fine for a start, but research is original work, and one cannot do original work just by doing what the supervisor tells you to do. Sooner or later you have to break free and be original, and the wise supervisor will do what he can to help. It is not easy to teach people how to be original.

If you are doing experimental work, there are numerous measurements that you could make. Which of them will give us new knowledge of nature? Existing theories may be a guide, but how do we know which is the best to use? We may have heard about the 'scientific method' and look for books telling us about it. There we read that the scientist first of all discovers facts, and then by a process called induction makes an hypothesis to try to account for them. From this hypothesis it is possible to deduce an infinite number of facts, including those we have already discovered. If these deduced facts do not agree with the observed facts, we modify the hypothesis or perhaps try another. If they do agree, the hypothesis is provisionally confirmed. We then look at some of the other facts deduced from the hypothesis and measure them. If they do not agree, we modify our hypothesis. If they do, we continue as before. It is always the mark of a good hypothesis that it agrees not only with the original facts used to construct it, but also predicts some new facts that are also confirmed.

At the same time, other scientists are studying different but related phenomena, and constructing hypotheses to account for the facts that they have discovered. Then by a higher type of induction a new hypothesis is found, from which both the previous hypotheses can be deduced. In this way successful hypotheses eventually become theories that really tell us something about how the world works. According to this account of the scientific method, there is a direct logical path from the observations and measurements to the theories.

This description of scientific research has some truth in it, but it is obviously not the whole story. If research were just a specified series of operations, it could just as well be carried out by a robot. Why then are some people so much better at it than others? Why do we find that groups of scientists like those led by Sommerfeld in Munich, Rutherford in Cambridge, Bohr in Copenhagen and Fermi in Rome were so outstandingly successful, whereas thousands of others achieved practically nothing?

It is just not true that there is a logical path from the observations and measurements to the theories. They are quite different, and to get from one to the other requires a leap of the imagination. By calling it induction, it is given a spurious appearance of being logical. Bacon thought that all one had to do was to amass facts until somehow the explanation would appear. But how does this happen? Sometimes we have no idea, other times we may think we know about what lies behind the appearances, so we are able to make a hypothesis about the nature and working of the world. We try it out in the way just described. It is a mysterious process; some people can do it and some cannot.

A great difficulty is that of deciding what to observe and measure. Most facts are useless for scientific purposes; how do we know which will be useful? We may say that if we are interested in a particular phenomenon, such as light, for example, we start by measuring everything we can about light. This again is easier said than done. We may not know even the existence of some aspect of light and so we cannot even begin to measure it. Research is pushing into the unknown and so, by definition, there are no maps to guide us. We may stumble on something new, but how do we know whether it is trivial or important?

Another way of doing science was due to Descartes. He was a great mathematician and emphasized that the world can be described mathematically. His method was to start with clear ideas that are self-evident and deduce everything from those ideas. He developed a theory of motion from such principles but never tested it to see if it was true. As it happened, his clear ideas were false, and so his conclusions about motion were all wrong.

The empiricism of Bacon and the rationalism of Descartes were thus both inadequate as accounts of the scientific method. Eventually, they were unified by Newton to give the essentials of the method that we all use today.

You might also ask yourself why you are never taught anything about the scientific method when you start to do research. At first sight this seems rather strange. If you propose to drive a car, you are carefully instructed for some months and then have to take a test. Why is there no similar preparation for the much more difficult task of doing scientific research? The reason is that it is not possible to learn just by reading instructions. In this respect it is just like learning how to drive a car; the only way is by doing it with the help of an instructor. Similarly, the only way to learn how to do research is to start doing it, under the guidance of someone who already knows.

Part of the trouble is that most of the accounts of the 'scientific method' are written by philosophers who have never been in a messy laboratory and tried for months to get a recalcitrant piece of apparatus to work. Have they ever struggled to find an explanation for some curious set of data that can be expressed mathematically? As Duhem once remarked, when infuriated by some philosophers who pontificated glibly about the scientific method: 'If you want to say something about science, first spend ten or fifteen years trying to find out about the natural world without any thought of philosophical applications'. So before you read a book about scientific method, ask how many papers the author has published in *Physical Review* or some similar journal. Rutherford had nothing but scorn for philosophers. He once remarked to the philosopher Samuel Alexander: 'When you come to think of it, Alexander, all that you have said and all that you have written during the last thirty years—what does it amount to? Hot air! Hot air!' This was not quite fair, and his disdain for philosophy was to cost him dearly later on.

The difficulty of understanding research is that it is impossible to do it just by reading articles in journals reporting the results of research. Anyone with actual experience of research knows that it is a chaotic medley of false starts, ideas that turn out to be wrong, apparatus that never works, calculations that become impossibly complicated ... We blunder from one failure to the next but maybe

*Corpus Christi College, Oxford OX1 4JF, U.K.
E-mail: p.hodgson1@physics.ox.ac.uk

sometimes light appears, fades and then becomes stronger until we realize that we have found out something new about the world. We sort out our ideas and gradually see that there is a path leading easily to our discovery. Then we write up the paper describing that smooth path, making it seem so easy and logical. We write it up as we would have liked it to happen, not how it actually happened, and we may even ourselves forget the tangled path we actually travelled. No editor would want to publish a truthful account of what actually happened; it would be far too long and no one would want to read it. The same may be said even more strongly for all the attempts at research that never achieved anything and were soon forgotten. Most of us have stacks of unpublishable would-be papers.

The accounts of science in textbooks are even further from what actually happened. Not only are the details omitted, but history is often re-written so as to lead the student as easily as possible to an understanding of the basic principles. The diagrams of the experimental apparatus are simplified and the reasoning re-written as a clear logical path from experiments to the final theory.

This has pedagogical advantages but is profoundly misleading as an account of how we actually reach a new truth. What actually happens is that many indications gradually accumulate, but each of them is inconclusive on its own. An example is provided by the gradual way scientists became convinced that matter is composed of atoms. The idea started with the ancient Greeks but they could not prove it to be true. It made sense of the regular structure of crystals, and the discovery by chemists that substances combine in definite proportions. At the end of the nineteenth century, the reality of atoms was still disputed by continental physicists. It was only gradually, mainly through the advances in nuclear physics, that the reality of atoms was universally accepted. Certainty is therefore seldom attained by a short logical path; instead it comes as the combination of many different results that can all be explained in the same way.

This way of attaining truth is not only characteristic of science; it applies also to the decisions of everyday life. It also helps us to understand why one person believes something and another does not. Belief depends on holding together in one's mind all the separate pointers, and without them belief is not possible.

An essential feature of scientific research is that theories must agree with experiment. Descartes put forward a theory of the motions of the planets, saying that they are carried round by vortices in the aether, but he did not formulate his idea mathematically or compare it with actual measurements. Newton had the idea that the same force that keeps the planets in their orbits causes apples to fall on earth. He did not leave it at that: he calculated the forces needed and found that they disagreed by about 20%, so he abandoned the idea. Later on, he heard about a new determination of the radius of the earth, a number that came into his calculation. He used this instead of the previous value and the results agreed, within the uncertainties of measurement. He then checked the idea against other measurements and they all agreed. Newton was doing physics; Descartes was not.

This union of fact and theory lies at the heart of physics. This has been emphasized by William Herschel:

> If we would hope to make any progress in an investigation of this delicate nature, we ought to avoid two opposite extremes, of which I can hardly say which is the most dangerous. If we indulge in fanciful imagination and build worlds of our own, we must not wonder at our going wide of the paths of truth and nature; but these will vanish like the Cartesian vortices, that soon gave way when better theories were offered. On the other hand, if we add observation to observation, without attempting to draw to only certain conclusions, but also conjectural views from them, we offend against the very end for which only observations ought to be made.[1]

The same point has been made by Ramón y Cajal: 'The hypothesis and the objective datum are linked together by a close aetiological relationship. To observe without thinking is as dangerous as to think without observing.'[2]

According to the 'scientific method', a theory is rejected if a measurement is made that contradicts it. This is not what happens in real research. For example, some early measurements on electrons gave results that were inconsistent with Einstein's theory but agreed with another theory. Einstein was unmoved by this: his theory was so beautiful that it must be true and the other so ugly that it could not be true. Eventually, a flaw was found in the experimental data and Einstein was vindicated. Similarly, he was not overly excited when Eddington verified his theory of gravitation, remarking that he knew that his theory was correct. While we say that a theory is not accepted until it agrees with experiment, it is equally true to say that an experiment is not accepted until it agrees with a theory. Perhaps it would be better to say that we know that we have attained at least some limited truth when there is a symbiotic resonance between experiment and theory.

An example of this is provided by the reception of the results of attempts to measure the velocity of the earth through the aether. The attempt by Michelson and Morley failed, and this was regarded as a great disappointment by some of the best physicists of the time. They believed in the existence of the aether and thought that there must be something wrong with the experimental data. Then Einstein's theory of relativity explained the result, and many other results as well, so that it became generally accepted. Years later, the very experienced physicist Dayton Miller repeated the Michelson-Morley experiment and obtained a small but definite effect. Once again, this result was not believed because it disagreed with a theory, but this time it was a different theory!

Scientific research is not a mechanical routine; it is more like a craft that can only be learned by being apprenticed to a master. Crease and Mann, when they were writing their book on science,[3] thought that it would be instructive to repeat Rutherford's famous experiment that showed the existence of the nucleus. So they asked Samuel Devons, who specialized in repeating historic experiments with the materials available to the pioneers, to show them how to do it. He simply laughed at their request, and told then that, though it was simple in principle, it was virtually impossible for two reasons. First of all, Rutherford used a radioactive source so strong that it would now be forbidden by the safety regulations. More subtly, he said that it was like asking a violin-maker to show them how to make a Stradivarius. Research is not just a matter of following instructions, it requires a mastery of numerous details of technique impossible to specify:

> Craft is a knowledge you have in your fingertips, little tricks you learn from doing things, and they don't work and you do them again. You have little setbacks and you think, how can I overcome them? And then you find a way. When you're pushing your equipment to the limits it's very easy to get spurious results; you've got to push what you know to the limit. If you don't, someone else is going to do it first.[3]

There were other groups trying to do the same thing as Rutherford at the same time, but they failed and no one ever hears of them. Rutherford could do it, and managed to pass his knowledge to many of his colleagues.

Millikan was one of many scientists trying to measure the charge on the electron. Other scientists got a range of values, and some even thought that it was a continuous variable. Millikan emphasized in his paper that all his measurements were made on single electrons and gave the same value, within the limits of experimental uncertainties. Examining his notebooks, it was found that he marked some examples: 'Beauty – publish', while others giving different results were discarded with the remark: 'Not an oil drop'. This was initially thought to be rather sharp practice, but properly understood it shows his genius as an experimentalist. The discordant results were indeed probably due to some other cause, such as dust.

Another remarkable experimentalist was J.J. Thomson. He was very clumsy and so his assistants had to keep him well away from his apparatus. He would make a brief sketch of what he wanted, and his assistants would set to work and make it.

When all was ready, he was allowed to switch it on. Usually it did not work, like most experimental apparatus. Then J.J. would sit and think about it and tell the assistants what needed to be done. He was removed until the work was finished and then the apparatus would usually work. He even managed to get the apparatus of other scientists working in the same way. He was a master of the experimental craft.

How did they manage to do such things? The answer to this question lies at the heart of the problem of how to do research. Somehow these scientists were able to see through the various experimental data to the underlying reality. For them, the measurements, the pointer readings are windows to the real. Knowing the reality, they were able to find new roads to show new aspects and with greater clarity. Indeed it often seemed that they did the experiments just to convince other people what they already knew themselves. Someone once suggested to Rutherford that alpha-particles are just mental constructs to unify our sense impressions and they do not really exist. Rutherford nearly exploded at this and burst out: 'Not exist! Not exist! I can SEE the little beggars in front of me!' And so of course he could.

Scientists inevitably describe their results in words, and then it is very easy for philosophers and others who have no direct experience of research to interpret the words in ways that are immediately repudiated by the scientists. Unlike the scientists, they do not see the reality behind the words. All scientists face this problem as students, when they start by reading textbooks and listening to lectures. Gradually they may come to see the reality, and this is strengthened if they do some research themselves. They succeed in different degrees, and some never learn to see. Nobody knows why.

It is difficult to describe this almost uncanny ability possessed by some scientists. Polanyi,[4] a scientist who thought deeply about science, described it by the phrase 'we know more than we can tell'. He also called it tacit knowledge. It is in some respects quite familiar, but the explanation is hidden in the subconscious. We all know how to ride a bicycle and to swim, but we could not describe it in such a way that someone listening could immediately ride and swim. We can all recall how we learned: we tried and failed, and suddenly we found we could do it. It is just the same as when we try to solve a mathematical problem. Another example given by Polanyi is the recognition of the face of a friend.[5] We can do this instantly, and yet we cannot describe that face so that someone else can recognize it with the same certainty.

In some respects this ability to see through the appearances to the real is like the ability of a radiographer to look at an X-ray photograph and see what it means for the body of the patient. Another example is that of the bubble-chamber physicist who is able to interpret the photograph of a jumble of tracks as the interactions of particular elementary particles. However, these skills can be learned, whereas we do not know how the scientist is able to acquire the ability to see through appearances to the reality beneath.

The ability to do research often depends on being able to separate the fruitful ideas from a vast mass of rubbish. Rutherford continually received reports of research from all over the world. Most of them he discarded, but when he saw something significant, he followed it up in every way possible.

Many scientists with heroic patience have devoted their lives to long series of measurements, and often the results are extremely useful. Thus Tycho Brahe spent about thirty years measuring the motions of the planets, and this provided the essential data for Kepler's epic breakthrough. Kepler spent about twenty years trying to determine the orbit of Mars, which, like everyone else, following Aristotle, he believed was a circle. He could fit it approximately but not to the accuracy of Brahe's measurements. He could have fudged it, but took the bold step of trying an ellipse. This fitted perfectly, within the accuracy of the measurements, and thus broke the stranglehold of Aristotle that had for so long prevented the development of physics. It is interesting to note that the accuracy of Brahe's measurements was just right for this discovery: if they had been less accurate, it would not have been possible to distinguish an ellipse from a circle, but if they had been more accurate an ellipse would not have fitted because of the perturbations due to the other planets. If Kepler had not done this work, it would probably never have been done, and who can say how modern physics would have arisen.

When engaged on routine measurements, it is essential to keep one's eyes open for the unexpected. This difficulty is that discrepant results happen quite often, and if one followed them all up there would be no time for anything else. Once again, this depends on the ability of the scientist. There are many examples of missed discoveries. The French astronomer Lalande measured the positions of thousands of stars. On 8 May 1795, he measured the position of a star and, when he checked it two days later, he found a different result. Some errors are bound to occur when one is making so many measurements, so he crossed out one of them. If he had followed this up, he would have realized that he had found a new planet, which was discovered in 1846 and called Neptune. Frederick Jervis-Smith was studying discharges of electricity through gases and noticed that photographic plates near the Crookes tube became fogged, so he told his assistant to store them further away. If he had followed up this observation, he would have discovered X-rays long before Roentgen. Hughes found that signals from a spark transmitter could be detected 500 yards away using a microphone, and claimed that they were electric waves in the air. This was dismissed by some eminent scientists, so he did not publish it until 1899. He had discovered radio waves seven years before Hertz. As Heraclitus remarked, 'if you do not expect the unexpected, you will not find truth'.

It is very difficult to know when to publish a new result. To be the first is essential, yet it is also important to be right. The line between glory and disgrace is a narrow one.

When making experiments or calculations to high precision, many small corrections have to be made. It is instructive to plot the measured values of the fundamental constants such as the charge on the electron or the velocity of light as a function of time. Quite often these show not only a steady convergence to more accurate values, but sudden jumps to new values outside the quoted experimental uncertainties of the previous determinations. The experimenters wanted their value to be consistent with previous work, but of course more accurate. As a result, they stopped making corrections when this was so. Eventually, some even more accurate measurements converged to a rather different value. These jumps occur when a scientist has the courage to publish a different result, and this is confirmed by subsequent measurements. Two examples of this phenomenon, one for the velocity of light and the other for the parameter that measures the violation of charge conjugation and parity in interactions of elementary particles, are given by Jeng.[6] This is called the bandwagon effect.

Jeng also gives several other examples of the effects of observer bias in physics research. Frequently what people see is strongly affected by what they expect to see, based on previous studies. Thus the early observations of the planet Saturn followed Galileo's sketches showing a large moon on either side, even though their telescopes were well able to show the rings. In his studies of optics, Newton used a prism that could easily show spectral lines, but he never reported them. It is strange that there are no European records of the supernova of 1054, although it was visible in daylight for several weeks and was observed in Japan and China. It is suggested that Europeans failed to see it because Aristotle said that there are no changes in the celestial realm. A remarkable example of seeing phenomena that are not there is provided by reports of N-rays in the early 1900s. About 300 papers on N-rays were published in France before it was finally concluded that they do not exist.

One of my research colleagues was scanning sensitive photographic emulsion using a high-power microscope and looking for events corresponding to the collision between a cosmic-ray particle and a nucleus in the emulsion. He noticed that there was quite often another event in the vicinity of the first. This was exciting because, if true, it implied that a particle with a high interaction cross-section was emitted from the first event and caused the second. No such particle was known, so had he discovered a new uncharged particle hitherto unknown to science?

To investigate this possibility, he examined many examples of pairs of events, and compared the number observed with that expected if the events were randomly spread over the emulsion. He did this as a function of the distance between the events, and this confirmed his earlier impression that there were far more closely associated events at small distances than would be expected by chance.

There seemed to be no explanation of this effect, and no means of investigating it further. At this point someone suggested that perhaps the emulsion in the vicinity of an event was searched more carefully than other parts of the emulsion. During normal searching, one moves the microscope continually across the emulsion, and only stops when one finds an event. When the microscope is stationary, it is more likely that one will see another event nearby. This was confirmed by careful re-examination, and so the effect was shown to be entirely spurious.

There are many examples of experimental measurements which give results that appear to be in excellent agreement with the theoretical predictions but which fail to be confirmed by more careful examination of the data. Thus Nichols and Hull measured the pressure due to light and found a value agreeing with Maxwell's theory to within 1%. A careful re-examination of their data produced a value 10% away from that calculated. Newton calculated the speed of sound in air and found a value agreeing with experiment to 0.1%. However, he assumed that the air is compressed isothermally instead of adiabatically, as is the case, and so should have found a value 15% too low. He managed to obtain the correct result by making a number of questionable corrections.

A similar problem is often encountered by a theoretician. In the course of his calculations he has to make a number of corrections to the principal result. For example, he may be calculating the orbit of a planet, so he begins by considering just the two-body system of the Sun and the planet. Then he has to make a series of corrections to take account of the presence of the other planets, the asteroids and so on. When does he stop? In this case the answer is relatively clear: when he finds that the effects of the smaller bodies are smaller than the accuracy he wishes to achieve. Many other cases are not so easy. I recall a distinguished physicist giving a lecture in which very many small correction terms were carefully calculated. After the lecture, someone asked him how he knew when to stop and received the reply: 'That's easy. I stop when the answer agrees with experiment!'

Theoreticians are often right, but not always, especially when they say that something cannot be done. They said that cyclotrons would not work above about 10 MeV because of relativistic effects; the experimentalists thereupon built the synchrocyclotron that went up to several hundred MeV. Pauli told Stueckelberg that his ideas about a massive particle to mediate the nuclear force were wrong, so he did not publish them. Afterwards, Yukawa independently got the same idea and published his theory of the pion. Later he reflected that the answer almost jumps at you, 'but my brain did not work so quickly. I had to take a wrong path first, before I could arrive at my destination. Those who explore the unknown world are like travellers without a map: the map is a result of the exploration. The position of their destination is not known to them, and the direct path that leads to it is not yet made.'

Bohr, saturated by the absurd Copenhagen interpretation of quantum mechanics, told Rutherford that he was wasting his time trying to find the structure of the nucleus, as it was just a structureless soup. Rutherford lacked the philosophical knowledge to refute him and gave up the search. Bohr similarly told Yamanouchi that his ideas on the nuclear shell model were wrong for the same reason; he did not publish and the credit went to Meyer and Jensen. Pauli criticized de Broglie's deterministic theory of quantum phenomena so severely that he abandoned it. This allowed the Copenhagen interpretation to be accepted, to the lasting confusion of physicists. Later on, Bell showed how Pauli's criticism could be answered.

One very fruitful way of carrying out theoretical research is to find some idea that everyone believes to be true, and then to show that it is wrong. Kepler did this for Aristotle's belief that the planets move in circles. Lee and Yang did this for the belief that parity is conserved. Another method is to find some beautiful equations and then see what phenomena they describe. Dirac did this most successfully. Einstein thought how God must have made the world. Such advice to would-be scientists is theoretically excellent, but in practice useless for ordinary mortals.

So, now we have discussed how not to do research, what we really want to know is how to do research. We have seen some of the things that must be avoided, and who we can trust to tell us. But how to see through appearances to reality cannot be taught. Some manage to achieve it, but even they cannot tell us how they did it. So the answer to our question how to do research is simply this: nobody knows.

I gratefully acknowledge the hospitality of the Department of Physics of Stellenbosch University and the continuing support of the Oppenheimer Foundation.

1. Hoskin M. (1959). *William Herschel: Pioneer of Sidereal Astronomy.* Sheed and Ward, London.
2. y Cajal R. (1989). *Recollections of My Life.* MIT Press, Cambridge, MA.
3. Crease R.P. and Mann C.C. (1986). *The Second Creation.* Macmillan, New York.
4. Polanyi M. (1958). *Personal Knowledge.* Routledge and Kegan Paul, London.
5. Polanyi M. (1966). *The Tacit Dimension.* Routledge and Kegan Paul, London.
6. Jeng M. (2006). A selected history of expectation bias in physics. *Am. J. Phys.* **74**: 578.